REMARKS

Claims 1-22 are pending in this application, of which claims 1, 4, 11, 15, and 22 are independent. In this Amendment, claims 1 and 2 have been amended. Care has been exercised to avoid the introduction of new matter. Support for the amendment of claims 1 and 2 can be found on, for example, page 6, line 15 to page 7, line 13 of the specification.

Claims 1-10 and 15-22 have been rejected under 35 U.S.C. §103(a) as being unpatentable over Shpiro et al. in view of Brandow et al., and further in view of Bruckert.

With respect to independent claim 1, the Examiner admitted that Shpiro et al. does not teach that the claimed storage device storing model phoneme array information including an array of phonemes and word boundaries of a sentence to be spoken by a learner; and the claimed word separation means for separating the sentence speech information on the basis of each word included in the sentence using the model phoneme array information. However, the Examiner asserted that Brandow et al. teaches the claimed separation means. The Examiner further admitted that the applied combination of Shpiro et al. and Brandow et al. does not teach the claimed storage device storing model phoneme array information including an array of phonemes and word boundaries of a sentence to be spoken by a learner. However, the Examiner asserted that Bruckert teaches the claimed storage device. On that basis, the Examiner concluded that the applied combination of Shpiro et al., Brandow et al., and Bruckert teaches the claimed foreign language learning device.

Applicants submit that Shpiro et al., Brandow et al., and Bruckert, either individually or in combination, do not disclose or suggest a foreign language learning device including all the

limitations recited in independent claim 1. Specifically, the applied combination does not teach, at a minimum, the following limitations:

> a storage device storing in advance a model sentence to be pronounced by a
> learner and model phoneme array information including an array of phonemes
> and word boundaries of said model sentence;
>
> word separation means for receiving sentence speech information, the sentence
> speech information corresponding to speech produced successively by the learner
> when the learner utters the model sentence including a plurality of words, to
> separate said sentence speech information into word speech information on the
> basis of each word included in said sentence using said model phoneme array
> information; ....

As cited above, claim 1 recites separating sentence speech information into word speech information. The sentence speech information is speech produced successively by a learner when the learner utters a <u>model sentence</u> including a plurality of words. To separate the sentence speech information into word speech information, the claimed subject matter uses model phoneme array information, stored in the storage device, including an array of phonemes and word boundaries <u>of the model sentence uttered by the learner</u>. A model sentence is prepared for the learner's practice and stored in the storage device in advance, and the model phoneme array information is prepared in relation to <u>that model sentence</u>. Claim 1 recites "storing in advance a model sentence to be pronounced by a learner and model phoneme array information including array of phonemes and word boundaries of <u>said model sentences</u>."

<u>Whether Shpiro et al. teaches using "model phoneme array information including an array of phonemes and word boundaries of the sentence."</u>

The Examiner asserted that Shpiro et al. teaches using "model phoneme array information including an array of phonemes and word boundaries of the sentence" to separate sentence speech information. To support this position, the Examiner relied upon column 5, lines 33-41; column 7, lines 16-29; and column 9, lines 24-39 of Shpiro et al.

Column 5, lines 33-41 of Shpiro et al. describes reference specimen player 100 operative to play reference audio specimens to student 110. The reference audio specimens are a multiplicity of phonemes, words and/or phrases. This paragraph is <u>irrelevant to the use of the model phoneme array information to separate sentence speech information</u>. The Examiner's reliance on this paragraph is technologically illogical.

Column 7, lines 16-29 describes how to segment students response specimen. According to Shpiro et al., a silence speech boundary is first identified as the point at which the energy increases to several times the back ground level and remains high, and then consonant-vowel boundaries are identified by identifying points at which the energy remains high but the dominant speech frequency increases to a range of about 100 to 200 Hz. These paragraphs do not teach using the claimed model phoneme array information including an array of phonemes and word boundaries of the sentence to be spoken by a learner. In addition, the Examiner did not provide any factual basis for supporting his position as to why these paragraphs teaches using the claimed model phoneme array information.

Column 9, lines 24-39, referring to Figs. 6 and 7, describes identifying the vowel "A" from the word "CAT." There is no suggestion in this portion relied upon by the Examiner that the claimed model phoneme array information includes an array of phonemes and word boundaries of the model sentence to be spoken by a learner.

As discussed in the April 27, 2007 response, the portions of Shpiro et al., upon which the Examiner relied, simply describe that reference audio specimens are a multiplicity of phonemes, words, and/or phrases (see column 5, lines 36-38), and a silence speech boundary and consonant-vowel boundaries are identified based on energy and frequency of a speech (see column 7, lines 16-29; and column 9, lines 24-39). Again, Shpiro et al. does not teach using the claimed model

13

phoneme array information including an array of phonemes and word boundaries of the model sentence to be spoken by a learner.

Applicants emphasize that the above assertion made by the Examiner is *illogical*. The Examiner asserted that Shpiro et al. teaches the claimed word separation means which utilizes the claimed model phoneme array information. To show that Shpiro et al. teaches the claimed word separation means, the Examiner has to show that Shpiro et al. has the claimed model phoneme array information including an array of phonemes and word boundaries of the sentence to be spoken. Without showing that Shpiro et al. has the claimed model phoneme array information, the Examiner is not allowed to conclude that Shpiro et al. teaches the claimed word separation means. The claimed word separation means requires the model phoneme array information including an array of phonemes and word boundaries of <u>the sentence to be spoken.</u>

Applicants specifically note that with respect to the rejections of claims 11-14, the Examiner admitted that Shpiro et al. "does not specifically teach that the storage device and word separation means includes *an array of phonemes and word boundaries of the sentence*" (see the second paragraph on page 10 of the Office Action) (emphasis added). *This admission is inconsistent with the above assertion regarding claim 1 that Shpiro et al. teaches the word separation means which uses the claimed model phoneme array including an array of phonemes and word boundaries of the sentence to be spoken.*

<u>Whether Brandow et al. teaches separating sentence speech information on the basis of each word including in a sentence using model phoneme array information.</u>

The Examiner states that Shpiro et al. does not teach that the word separation means separates sentence speech information on the basis of each word included in a sentence using model phoneme array information. Brandow et al. does not cure the deficiencies of Shpiro et al.

for the reasons set forth below. The Examiner has to show that Brandow et al. uses the claimed model phoneme array information.

Brandow et al. pertains to a speech recognition correction system for correcting text data generated from a speech recognition system. The Examiner cited column 1, lines 21-35 (reproduced below (emphasis added)), and asserted that Brandow et al. teaches separating sentence speech information (segment speech) on the basis of each word included in a sentence using model phoneme array information.

> Speech recognition is the process of converting an acoustic signal, captured by a microphone or a telephone, to a set of words. In a typical speech recognition system, the acoustic signal is converted into a digitized speech signal and then segmented into a set of speech segments. Each set of speech segments contains useful measurements or features known as phonemes. Phonemes are the smallest sound units of which words are composed of. The phonemes are then represented by using a phonetic language model such as a 2-phoneme or 3-phoneme hidden Markov model (HMM). The HMM captures and represents patterns of variation of the phonemes into phoneme groups. <u>The phoneme groups are then applied to a language model such as a 2-gram or 3-gram HMM, which is used to recognize the most probable words for each group and then transcribe the words.</u>

The cited portion simply describes the process of converting an acoustic signal, captured by a microphone or telephone, to a set of words by use of language model such as a hidden Markov model (HMM). The HMM is irrelevant to the claimed model phoneme array information. The claimed model phoneme array information includes array of phonemes and word boundaries of the model sentence which is stored in the storage device and is spoken by a learner. Brandow et al. does not teach that the HMM uses such information.   It is, therefore, apparent that *a person skilled in the art having common sense does not consider that Brandow et al. teaches using the claimed model phoneme array, and he/she does not utilize the teachings of Brandow et al. to cure the deficiencies of Shpiro et al. in order to arrive at the claimed subject matter.*

Based on the above, Brandow et al. is silent on the claimed model phoneme array information which includes array of phonemes and word boundaries of sentences to be spoken by a learner, and which is used to separate sentence speech information. The deficiencies of Shpiro et al. are not cured by Brandow et al.

Whether Bruckert teaches a storage device storing model phoneme array information including an array of phonemes and word boundaries of a sentence to be spoken by a learner.

According to the first to third paragraphs on page 5 of the Office Action, Bruckert was cited because the applied combination of Shpiro et al. and Brandow et al. does not teach a storage device storing model phoneme array information including an array of phonemes and word boundaries of the sentence to be spoken by a learner. However, as set forth above, *the applied combination of Shpiro et al. and Brandow et al. does not teach using the claimed model phoneme array information including an array of phonemes and word boundaries of the sentence to be spoken by a learner, in order to separate the sentence speech information.* Accordingly, the Examiner's assertion that Bruckert teaches the claimed storage means for storing model phoneme array information is not viable, and Bruckert does not cure the deficiencies of the applied combination of Shpiro et al. and Brandow et al.

Bruckert pertains to a method for generating synthetic speech using detection of natural timing boundaries in words to be spoken by the synthetic speech system to produce natural timing intervals (see the Abstract). Accordingly, it is apparent that Bruckert does not teach, among other things, the word separation means for separating sentence speech information into word speech information on the basis of each word included in the sentence stored in the storage

device using the model phoneme array information, as claimed. Therefore, Bruckert does not cure the deficiencies of the applied combination of Shpiro et al. and Brandow et al.

Not all the claimed limitations are taught by the cited references.

Based on the foregoing, Shpiro et al., Brandow et al., and Bruckert, either individually or in combination, do not disclose or suggest a foreign language learning device including all the limitations recited in independent claim 1 within the meaning of 35 U.S.C. §103. The above discussion is applicable to independent claims 4, 15, and 22. Dependent claims 2, 3, 5-10, and 16-21 are also patentably distinguishable over Shpiro et al., Brandow et al., and Bruckert at least because these claims include all the limitations recited in independent claims 1, 4, and 15, respectively. Applicants, therefore, respectfully solicit withdrawal of the rejection of claims 1-10 and 15-22 under 35 U.S.C. §103(a), and favorable consideration thereof.

Claims 11-14 have been rejected under 35 U.S.C. §103(a) as being unpatentable over Shpiro et al. in view of Acero and further in view of Bruckert.

With respect to independent claim 11, the Examiner admitted that Shpiro et al. does not teach that "the storage device and word separation means includes an array of phonemes and word boundaries of the sentence." However, the Examiner asserted that Acero teaches the word separation means (figure 4, element 294). The Examiner further admitted that the applied combination of Shpiro et al. and Acero does not teach a storage device storing model phoneme array information including an array of phonemes and word boundaries of the sentence to be spoken by a leaner. However, the Examiner asserted that Bruckert teaches the storage device. On that basis, the Examiner concluded that the applied combination of Shpiro et al., Acero, and

Bruckert teaches the claimed foreign language learning device. This rejection is respectfully traversed.

Applicants submit that Shpiro et al., Acero, and Bruckert, either individually or in combination, do not teach a foreign language learning device including all the limitations recited in independent claim 11. Specifically, the applied combination does not teach, among other things, the following limitations of claim 11 (emphasis added):

> storage means for storing a model sentence to be pronounced by a learner and model phoneme array information including an array of phonemes and word boundaries <u>corresponding to said model sentence</u>; ...

> word separation means for receiving sentence speech information corresponding to a sentence pronounced by said learner to separate the sentence speech information into word speech information <u>on the basis of each word included in said sentence</u>....

It is submitted that Shpiro et al. is silent on the claimed model phoneme array information which includes array of phonemes and word boundaries of sentences to be spoken by a learner, and which is used to separate sentence speech information. Applicants incorporate herein the arguments regarding Shpiro et al. made in responding to the rejection of claim 1 under 35 U.S.C. § 103 for obviousness predicated upon Shpiro et al., Brandow et al., and Bruckert.

The secondary reference, Acero, pertains to a <u>speech *synthesizing* system</u>, and hence, does not disclose a system recognizing learner's utterances by using text data. The reference does <u>not have any information on sentences to be recognized beforehand</u> (see the storage means in claim 11), and does not teach any means for separating sentence the received. *Acero does not disclose or suggest any means which meets the requirements of claim 11 reciting the word separation means. Applicants respectfully request the Examiner to reconsider Applicants argument discussed below.*

The Examiner identified parser 294 in Fig. 4 of Acero as the claimed word separation means. However, parser 294 is not provided to separate sentence speech information into word speech information. According to Acero, parser 294 parses training text 282 into sub-word units or states (column 5, lines 14-16). Text 282 is not the claimed speech information, but a text itself because Acero describes "input speech 280 is generated by a speaker while reading text 282" (column 4, line 40-42), i.e., <u>Acero pertains to a speech synthesizing system</u>. Furthermore, Acero does <u>not</u> teach <u>using text 282 to separate sentence speech information into word speech information based on words included in the text</u>. In contrast, claim 11 recites "<u>receiving sentence speech information corresponding to a sentence pronounced by said learner</u> to separate the sentence speech information into word speech information on the basis of each word included in said sentence."

The third reference, Bruckert, pertains to a method for <u>generating synthetic speech</u> using detection of natural timing boundaries in words to be spoken by the synthetic speech system to produce natural timing intervals (see the Abstract). Accordingly, it is apparent that Bruckert does not teach, among other things, the claimed word separation means, and does not cure the deficiencies of the applied combination of Shpiro et al. and Acero.

Based on the foregoing, Shpiro et al., Acero, and Bruckert, either individually or in combination, do not teach a foreign language learning device including all the limitations recited in independent claim 11 within the meaning of 35 U.S.C. §103. Dependent claims 12-14 are also patentably distinguishable over Shpiro et al., Acero, and Bruckert at least because these claims include all the limitations recited in independent claims 11. Applicants, therefore, respectfully solicit withdrawal of the rejection of the claims and favorable consideration thereof.
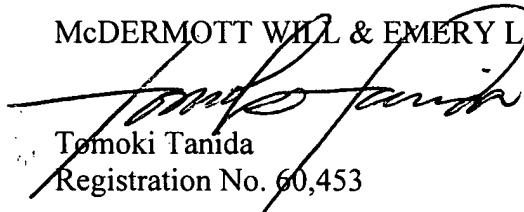
Conclusion

It should, therefore, be apparent that the imposed rejections have been overcome and that all pending claims are in condition for immediate allowance. Favorable consideration is, therefore, respectfully solicited.

To the extent necessary, a petition for an extension of time under 37 C.F.R. 1.136 is hereby made. Please charge any shortage in fees due in connection with the filing of this paper, including extension of time fees, to Deposit Account 500417 and please credit any excess fees to such deposit account.

Respectfully submitted,

McDERMOTT WILL & EMERY LLP

Tomoki Tanida
Registration No. 60,453

600 13th Street, N.W.
Washington, DC 20005-3096
Phone: 202.756.8000 SAB:TT
Facsimile: 202.756.8087
Date: December 26, 2007

Please recognize our Customer No. 20277 as our correspondence address.

WDC99 1505561-1.057454.0235

20